

# Challenges in Mathematical Cognition: A Collaboratively-Derived Research Agenda

## Author Names

Author 1: Lara Alcock\*

Author 2: Daniel Ansari

Author 3: Sophie Batchelor

Author 4: Marie-Josée Bisson

Author 5: Bert De Smedt

Author 6: Camilla Gilmore

Author 7: Silke Melanie Göbel

Author 8: Minna Hannula-Sormunen

Author 9: Jeremy Hodgen

Author 10: Matthew Inglis

Author 11: Ian Jones

Author 12: Michèle Mazzocco

Author 13: Nicole McNeil

Author 14: Michael Schneider

Author 15: Victoria Simms

Author 16: Keith Weber

## **Author Affiliations**

Author 1: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 2: Numerical Cognition Laboratory, Department of Psychology & Brain and Mind Institute, The University of Western Ontario, London, Canada

Author 3: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 4: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 5: Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Author 6: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 7: Department of Psychology, University of York, York, UK

Author 8: Department of Teacher Education, University of Turku, Turku, Finland

Author 9: School of Education, University of Nottingham, Nottingham, UK

Author 10: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 11: Mathematics Education Centre, Loughborough University, Loughborough, UK

Author 12: Institute of Child Development, University of Minnesota, Minneapolis, USA

Author 13: Department of Psychology, University of Notre Dame, Notre Dame, USA

Author 14: Educational Psychology, Universität Trier, Trier, Germany

Author 15: Psychology Research Institute, Ulster University, Belfast, UK

Author 16: Graduate School of Education, Rutgers, The State University of New Jersey, New Brunswick, USA

## **Corresponding Author's Contact Address**

Mathematics Education Centre

Loughborough University

Loughborough

LE11 3TU

UK

ACCEPTED MANUSCRIPT

## **Abstract**

This paper reports on a collaborative exercise designed to generate a coherent agenda for research on mathematical cognition. Following an established method, the exercise brought together 16 mathematical cognition researchers from across the fields of mathematics education, psychology and neuroscience. These participants engaged in a process in which they generated an initial list of research questions with the potential to significantly advance understanding of mathematical cognition, winnowed this list to a smaller set of priority questions, and refined the eventual questions to meet criteria related to clarity, specificity and practicability. The resulting list comprises 26 questions divided into six broad topic areas: elucidating the nature of mathematical thinking, mapping predictors and processes of competence development, charting developmental trajectories and their interactions, fostering conceptual understanding and procedural skill, designing effective interventions, and developing valid and reliable measures. In presenting these questions in this paper, we intend to support greater coherence in both investigation and reporting, to build a stronger base of information for consideration by policymakers, and to encourage researchers to take a consilient approach to addressing important challenges in mathematical cognition.

## **Keywords**

mathematical cognition, numerical cognition, research agenda, mathematics education

# Introduction

Research on mathematical cognition has made substantial advances in recent years, in areas that span diverse theoretical frameworks both within and across disciplines. For instance, neuroscience research has revealed potential mechanisms contributing to the representation of number and related concepts in the brain (De Smedt, Noël, Gilmore, & Ansari, 2013). Behavioural research has identified distinctive characteristics in the mathematical thinking of young children with atypical development (Geary, 2010; Kaufmann et al., 2013; Dennis, Berch, & Mazzocco, 2009), individual differences in typical mathematics learning (Geary, 2011; Raghubar, Barnes, & Hecht, 2010), and the nature and variety of mathematical expertise (Weber, Inglis, & Mejía-Ramos, 2014). Studies on the effects of the environment, including home (LeFevre et al., 2009) school (Clements & Sarama, 2004; Beilock, Gunderson, Ramirez & Levine, 2010), and culture (e.g., Jones, Inglis, Gilmore, & Dowens, 2012), have also influenced theoretical conceptualisations of mathematical learning and contributed to the design of interventions to support that learning (Cohen Kadosh, Dowker, Heine, Kaufmann, & Kucian, 2013).

Collectively, this work could have substantial international policy implications, and indeed specific research programmes have led to large-scale trials of teaching practices and materials development (Torgerson, Wiggins, Torgerson, Ainsworth, & Hewitt, 2013). However, the potential of the field as a whole has yet to be fully realised, perhaps in part due to its diverse nature. This diversity affords the benefits that accompany multiple perspectives, but also generates obstacles resulting from discrepant methodologies. So, while mathematical cognition researchers share central concerns – the desire to understand and support mathematical thinking and learning at all educational levels – they have different priorities, methods and conceptions of central theoretical constructs, and their work is published across a diverse range of outlets. Arguably this is an appropriate time to draw together current thinking from these fields and to articulate a mutually understood view of research priorities that can inform and shape future research in mathematical cognition.

This paper is designed to do that. It reports the results of an exercise that brought together 16 researchers from six countries with expertise in mathematics education, psychology and cognitive neuroscience, and that sought to identify key open questions in mathematical cognition. The questions were generated using a collaborative process modelled on previous exercises in other fields (Sutherland et al., 2006, 2012; Sutherland, Fleishman, Mascia, Pretty, & Rudd, 2011), with adaptations appropriate to the comparatively focused research topic and small number of participants. We followed a six-stage process to generate an initial list of questions, select questions that participants collectively viewed as most important, refine these so as to meet a list of inclusion criteria, and organise them in a form conducive to the discussion and planning of new and ongoing research programmes (for details see the Method section). From these questions, practical suggestions evolved regarding how to maximize the field's ability to share both work and findings across research laboratories and disciplines.

The final question list, as presented in the Results section, is narrow in the sense that it is specifically focused on research related to the cognitive aspects of human mathematical thinking and learning. But the list is broad in that it reflects priorities ranging from issues of representation in the brain, through the mathematical thinking and learning of very young and older children, up to and including the practices and reasoning of expert mathematicians. We intend the questions to support greater coherence in both investigation and reporting of issues in mathematical cognition, and thus to build a stronger base of information for consideration by researchers and policymakers. We hope that the questions will inspire researchers in the relevant fields to likewise address big challenges in mathematical cognition and, in the process, to develop methods that promote interconnected theories.

## **Background**

Researchers with interests in mathematical cognition represent a variety of disciplines.

Those working in mathematics education have often trained initially as mathematicians or teachers; some have moved from an undergraduate or masters experience in mathematics straight through to a doctorate in education, but many have moved into educational research later in their careers after teaching in schools, working to prepare teachers, or working as mathematicians. They typically publish in outlets such as the *Journal for Research in Mathematics Education*<sup>1</sup>, *Educational Studies in Mathematics* and the *Journal of Mathematical Behavior* and share work at conferences such as that of the *International Group for the Psychology of Mathematics Education*. Much work in all of these outlets comprises detailed qualitative studies (including case studies) on students' understandings of specific mathematical concepts, from those in early number right up through undergraduate study and into work on the reasoning of mathematical experts. Other work takes the form of design research (Prediger, Gravemeijer & Confrey, 2015), and experimental studies are comparatively rare (Alcock, Gilmore, & Inglis, 2013).

Those working in psychology have usually trained initially in developmental, experimental, or cognitive psychology; some pursue research in mathematical cognition due to an interest in mathematics per se, but others are interested in more general cognitive functioning and use mathematical ideas primarily as a vehicle to study this. These researchers typically publish in journals including *Cognition*, *Child Development*, *Developmental Science*, *Journal of Experimental Child Psychology*, *Journal of Educational Psychology* and *Learning and Instruction*, and share work at conferences such as that of the *European Association for Research on Learning and Instruction* and the *Society for Research in Child Development*. Unlike those with backgrounds in education, they frequently conduct experimental or quasi-experimental studies, but the mathematical concepts involved are often based in simple arithmetic and its precursors.

Those working in cognitive neuroscience have often trained initially in experimental or cognitive psychology, cognitive science or neuroscience. Most have moved into developmental cognitive neuroscience after doctoral training in adult cognitive neuroscience

and/or developmental disorders. These researchers typically publish in journals such as the *Journal of Cognitive Neuroscience*, *Neuropsychologia* and *Trends in Education and Neuroscience*, and share work at conferences such as the *Annual Meeting of the Cognitive Neuroscience Society* and *Conference of the European Society for Cognitive Psychology*. Typical studies in this area include measures of brain activity during number processing and arithmetic; often the work relates these to behavioural measures. Central research questions involve investigating the brain correlates of number processing and arithmetic, how these change with development, skill acquisition and training, and their role in dyscalculia.

Crucially for the development of research in mathematical cognition, there has traditionally been little communication between researchers working in these areas; cross-citations (with education in particular) are comparatively rare. If this is a fault then it is a common and predictable one: researchers across all disciplines publish their work primarily in those journals that most closely reflect their influences and methods, so it is natural that citation silos should evolve. But this clearly limits an interdisciplinary endeavour: insights from one area might well inform, accelerate, or supplant theoretical developments in another, and methodological diversity should be embraced as a means of developing both different types of insight and individual researcher skills. For these reasons we sought to assemble researchers with a diverse range of interests in human mathematical cognition, and to implement a process that would focus their attention on their common goals, clarify similarities and differences in theories and methods, and facilitate future communication. In communicating the outcomes of this process in the form of a paper, we aim to further these goals by stimulating a similar debate across the broader mathematical cognition community. We believe that the time is right for this because a collective desire for communication is evident in the establishment of interdisciplinary societies (e.g., the International Mind, Brain, and Education Society) and journals like this one, and because one way to maximise the gains from structured cross-talk is to focus on the big questions we would all like to see answered. We return to this point in the Discussion section of this paper.



## Method

We modelled our method on that described in Sutherland et al. (2012), who pursued a Delphi-like process (Hsu & Sanford, 2007) to identify a list of unanswered questions on the relationship between science and policy; Sutherland has also been involved in similar exercises in ecology and conservation (e.g., Sutherland et al., 2006)<sup>ii</sup>. We were conscious throughout that, as Sutherland et al. (2011, p.246) note “[a]ny priority-setting exercise is the product of the people who participate. The results can be influenced by the interests present. Furthermore, individuals may have agendas. However, a diverse and moderately large group, clear criteria, and a democratic process all help reduce the impact of any individual.” In this section we explain how we addressed these issues.

Compared with fields like science policy, our research domain is small. But we aimed nevertheless to act in the spirit of Sutherland et al.’s (2012) approach, seeking participants who would represent a range of disciplines and who would ideally have experience or at least knowledge of the theories and methods typically in use in two or more areas. Because resources were limited and because we were interested in drawing together a community of researchers whose work could be expected to influence the field for an extended time, we sought particularly to invite early- to mid-career academics with research programs that were established enough to have a high probability of making substantial further contributions in the coming decades. With these goals in mind, we invited 16 scholars from the disciplines of mathematics education, psychology and cognitive neuroscience to take part (the 16 are the authors of this paper); 15 attended as delegates at the Royal Society-funded conference *Grand Challenges in Mathematical Cognition*<sup>iii</sup> and one participated in all the non-conference stages and had a particular role in organising the questions for presentation. Table 1 lists the participants’ countries of affiliation, main research areas, and specific illustrative research foci.

Table 1: Participants' Main Research Areas and Specific Illustrative Research Foci

Participant	Country	Edu	Psych	Neuro	Illustrative Research Foci
Lara Alcock	UK	X			Undergraduate proof comprehension Expert and novice mathematical reading Learning interventions in undergraduate mathematics
Daniel Ansari	Canada		X	X	Basic number processing and individual differences Cognitive and neuronal correlates of developmental dyscalculia Symbolic and non-symbolic magnitudes at behaviour and brain levels
Sophie Batchelor	UK	X	X		Children's magnitude representations and counting skills Spontaneous focusing on numerosity and arithmetic skills Mathematics anxiety in primary school children
Marie-Josée Bisson	UK	X	X		Measuring conceptual understanding in mathematics Using comparative judgement in randomised controlled trials Incidental learning

Participant	Country	Edu	Psych	Neuro	Illustrative Research Foci
Bert De Smedt	Belgium	X	X	X	Basic number processing and individual differences Dyscalculia; mathematics learning difficulties Arithmetic (strategy use)
Camilla Gilmore	UK		X		Early numerical development Executive functions and mathematics Educational outcomes of prematurity
Silke Melanie Göbel	UK		X	X	Longitudinal predictors of arithmetic growth in primary school Number-space associations Language influences on number processing
Minna Hannula-Sormunen	Finland	X	X		Early numerical development Spontaneous focusing on numerosity and quantitative relations Learning environments for early numeracy and flexible arithmetic skills
Jeremy Hodgen	UK	X			Learning interventions in school mathematics

Participant	Country	Edu	Psych	Neuro	Illustrative Research Foci
Matthew Inglis	UK	X			Educational standards over time
					Post-compulsory mathematics participation
					Mathematical problem solving and reasoning development
					Expert mathematical practice
					The Approximate Number System
Ian Jones	UK	X			Assessing mathematical understanding
					Children's understanding of equivalence
					Students' understanding and progression in school mathematics
Michèle Mazzocco	USA	X	X		Individual differences in mathematical cognition
					The development of early number concepts
					Mathematics learning difficulties; dyscalculia
Nicole McNeil	USA		X		Optimising input to facilitate children's understandings of key mathematical concepts
					Benefits and disadvantages of concrete and abstract representations

Participant	Country	Edu	Psych	Neuro	Illustrative Research Foci
Michael Schneider	Germany		X		Predicting individual differences in early mathematical achievement
					Development of conceptual and procedural knowledge
					Conceptual change
					Numerical magnitude processing
Victoria Simms	UK		X		Disordered mathematical cognition
					Estimation skills and their impact on mathematical development
					General cognitive factors influencing mathematical cognition
Keith Weber	USA	X			Mathematical reasoning
					Undergraduate mathematics education
					Construction and comprehension of mathematical proofs

*Note.* Edu = Education; Psych = Psychology; Neuro = Neuroscience

With the participants in place, the questions that form the core of this paper were generated in six stages: Stages 1 and 2 were completed online, Stages 3 to 5 occurred at the conference and Stage 6 took place after the meeting.

In Stage 1, all participants were invited to draw up a list of research questions for which answers would significantly advance our understanding of mathematical cognition. They were encouraged to draw on their own work and that of their broader discipline, as well as their knowledge of classic problems or recent developments; they were also encouraged to solicit input from colleagues and contacts. The team of organisers collated these questions, amalgamated or removed duplicates, and organised the resulting 155 questions into themes. In Stage 2, these questions were anonymised and made available in an online survey via which all participants rated all questions on a scale of 1 (low) to 10 (high) according to their importance for advancing the field. The original questions and original and summarised rankings are available at [https://figshare.com/articles/S1\\_Dataset\\_xlsx/1299358](https://figshare.com/articles/S1_Dataset_xlsx/1299358). After the survey, the 63 questions with a median score of 8 or higher were retained for the conference stages.

At the conference, Stages 3, 4 and 5 of the question winnowing and refinement process were completed. Delegates were asked to base their question selection and refinement on criteria drawn from Sutherland et al. (2011) stating that the eventual questions should:

- address an important gap in knowledge;
- be formulated specifically (not as a general topic area);
- be clear, where appropriate, about specific interventions and outcome measures;
- be answerable specifically (not by 'it all depends');
- have a factual answer that does not depend on value judgements;
- be answerable through a realistic research design;
- be of a scope that could reasonably be addressed by a research team.

For Stage 3, delegates were assigned to subgroups of three or four participants representing the variety of research backgrounds present across the assembly of 16. Sufficient time was provided for members of the mixed groups to discuss all the questions, noting ambiguities and suggesting initial rephrasings. For Stage 4, delegates were assigned to topic groups composed of four or five participants, where the topics were named *numerical representations and developmental dyscalculia*, *cognitive factors beyond basic representations*, and *beyond basic numeracy*. The topic groups were tasked with reducing their respective question lists by half, and with rewriting questions where appropriate to reduce ambiguity or duplication and to ensure that the questions met the aforementioned inclusion criteria. For Stage 5, all 15 participants reconvened and reviewed the remaining lists. At this stage further refinements to wording were suggested and questions that appeared essentially equivalent across two of the lists were combined. We note that the discussions were extremely constructive throughout, and we found that our conceptions of key constructs and issues in learning were remarkably similar. Resolving differences was therefore largely a matter of settling on terminology that did not evoke unintended ideas or assumptions for any one group.

Stage 6 took place after the conference. We were conscious that the assignment of delegates to topic groups necessarily stratified the questions by the educational stage of their typical research participants. Thus the question lists were merged and were sent to the non-attending participant for sorting into smaller groups without this restriction. The outcome grouped the questions in a way that focused on conceptual relationships rather than educational stage; these groups of questions appear in the Results section.

## **Results: Research Questions**

We present the results of the exercise in two sections. In the first section, we list the groups of questions that emerged from Stage 3 (as described above); each group (labeled A-F) is accompanied by a brief commentary. We present the questions in generic forms, rather than

from the perspectives of any specific theoretical frameworks, because each might profitably be addressed using multiple frameworks (as is reflected in the field). Where there have been recent developments in competing theories, we endeavour to draw attention to these. In the second section, we take two subgroups of questions – those on the nature of mathematical thinking and those on developmental trajectories and their interactions – and offer more detailed critical review of recent developments in these areas. We take this approach because it enables us not only to illustrate the range of work taking place across research in mathematical cognition but also to discuss the relative maturity of different parts of the field, thereby highlighting the shifting theoretical and methodological demands that arise as research areas develop.

## **A. Elucidating the Nature of Mathematical Thinking**

1. Do infants really have a 'sense of number' or are they merely sensitive to quantitative dimensions?
2. How do feelings of correctness or doubt arise when people are doing mathematics, and how do they influence mathematical reasoning and response to instruction?
3. What strategies do experts and students use when evaluating whether a mathematical assertion is true or false?
4. What mathematical cognition do experts use in their practice and what methods can researchers use to provide an accurate portrayal of this?

These questions reflect the ambitious long-term goal of understanding mathematical cognition from infancy to mature expertise. Collectively they address the contribution of nativist theories in proposing foundational “number sense” capacities seen in nearly all human infants (e.g., Feigenson, Dehaene, & Spelke, 2004; Geary, 2007), the mature state of refined formal cognitive capabilities associated with highly successful mathematicians (Weber, Inglis, & Mejía-Ramos, 2014), and the as-yet unexplained mechanisms that link the two. Although individual research studies necessarily focus on a relatively narrow part of this



spectrum, participants in the exercise recognised the importance of working towards an empirically evidenced understanding of mathematical cognition that accurately captures effective mathematical reasoning, that provides reliable ways of diagnosing atypical development, and that is specific enough to be leveraged for the design of interventions. In the Critical Review section to follow, we discuss the methodological challenges associated with studying expertise in particular.

## **B. Mapping Predictors and Processes of Competence Development**

5. How do children learn the meaning of non-iconic representations of number?
6. What are the key mathematical concepts and skills that children should have in place prior to the start of compulsory education?
7. Do causal mechanisms underlie the correlational evidence between domain-specific foundational competencies (e.g., the Approximate Number System, (ANS)) and mathematics performance? What are reliable early and later longitudinal predictors of the development of number skills, arithmetic and other aspects of mathematics?

These might include:

- a. domain specific factors;
  - b. domain general factors (working memory, executive functions, language, spatial skills);
  - c. non-cognitive factors (motivation, anxiety);
  - d. environmental factors (home, school, outside of school learning, early education).
8. What are the interactions amongst these predictors?
  9. What are the directions of associations?

These questions are motivated by concerns about the large proportion of individuals of all ages whose achievement in mathematics is considered inadequate relative to standards or criteria specified by governments or related agencies. Compared with the questions in Set A, they are more specific in several ways. First, they explicitly acknowledge that both typical

and atypical mathematical development might be influenced by a wide range of environmental and affective factors, as well as by domain-specific and domain-general cognitive competences. As such, it is necessary to consider the full range of influences (and the ways in which we measure these) because such factors might interact in complex ways. Moreover, the nature of these interactions might vary across periods of development. This issue becomes more salient as apparent inconsistencies across research studies come to light, as illustrated by studies of the Approximate Number System (ANS) and formal numerical skills (De Smedt, Noël, Gilmore & Ansari, 2013; Gilmore et al., 2013). The time is ripe for studies designed to disentangle the effects of different predictors and to map patterns of development, and such studies are beginning to be reported (Fuchs et al., 2010; Göbel, Watson, Lervåg, & Hulme, 2014; Klibanoff, Levine, Huttenlocher, Vasilyeva & Hedges, 2006; Mazzocco, Feigenson & Halberda, 2011).

### **C. Charting Developmental Trajectories and Their Interactions**

10. Are there alternative cognitive pathways by which one can be a successful mathematics student? If yes, what does this imply for typical research methods used by mathematical cognition researchers?
11. How are different mathematical skills (including representing number, counting, performing arithmetic, using fractions) and their developmental trajectories related to each other?
12. Is developmental dyscalculia qualitatively different from arithmetic performance at the lower end of the normal distribution?
13. As we build causal models of pathways to mathematics learning and thinking, what models emerge for compensatory mechanisms – cognitive skills or abilities that allow individuals to overcome shortcomings in foundational skills, and that contribute to patterns of behaviors?

The questions in Set B acknowledge the continuum of typical and atypical development, but there is no reason to presume that pathways to mathematical understanding are

homogeneous among all children whose path is considered either “typical” or “atypical”; in both scenarios, individual differences are likely (Geary, 2011; Kaufmann et al., 2013, respectively). The questions in Set C thus consider how research might capture variations in mathematical learning. They are particularly pertinent for researchers considering cognition about mathematical concepts beyond basic arithmetic because, earlier points notwithstanding, early number and basic arithmetic have been comparatively well studied. This is partly because arithmetic concepts provide a clean context for setting and evaluating performance on a restricted set of problems with definite right and wrong answers, meaning that these concepts provide a sensible setting for investigating more general cognitive processes (e.g., strategy use, for which see Siegler, 1988; Wu et al., 2008; gesture, for which see Goldin-Meadow, Wagner Cook & Mitchell, 2009). They are also amenable to investigation in the sense that research has identified a fairly restricted number of strategies that a student might use to solve, for instance, a single-digit addition problem. But mathematics as a whole has a hierarchical structure that builds rapidly through many levels of abstraction. At every level, it might be possible to explain learning not only in terms of what students lack – mature working memory systems, proficiency with basic arithmetic facts – but also in terms of what they already know: previous knowledge can interfere with the learning of higher-level ideas (e.g., McNeil, 2014). Researchers have developed pockets of knowledge about how this might play out (e.g., the natural number bias as discussed by Van Hoof, Lijnen, Verschaffel, & Van Dooren, 2013). But knowledge development is currently understood only in single steps in which a particular existing conception is known to influence interpretations of a new concept – we know very little about longitudinal development in mathematical cognition, and the questions in this section reflect this.

## **D. Fostering Conceptual Understanding and Procedural Skill**

14. How can educators effectively capitalise on informal mathematical knowledge when teaching mathematical concepts?

15. How can instruction help students to identify normatively-valued similarities between mathematical domains, problems, representations, and situations?
16. How can we help children develop fluency with basic facts and skills while still promoting understanding of the underlying concepts?
17. How can we teach/learn so that useful mathematical knowledge gets automated, while also promoting understanding of the underlying concepts?
18. For specific concepts, which students profit from different sequences of exposure to more and less concrete representations?

Mathematical thinking does not emerge in a vacuum. By definition, emergence of formal mathematics requires instructional input, but what form should this instruction take? The questions in Set D sit at the heart of a distinction that causes consternation and controversy. Students at all levels must develop procedural fluency so that they can efficiently execute standard manipulations, and conceptual understanding so that they can identify and work effectively with mathematical structures that arise in a variety of situations (NRC, 2001). The fundamental distinction between the two is conceptualised in various guises and terminologies in both education (Skemp, 1976; Star, 2005) and psychology (Rittle-Johnson & Schneider, 2015), and the apparent tension manifests itself most dramatically in the so-called “math wars” (Schoenfeld, 2004), which contrast “back to basics” movements advocating drill and practice – intended to develop factual knowledge and procedural fluency – with programmes designed to foster conceptual understanding by providing mathematical learning situations that are personally meaningful to the student (Gravemeijer, 1994; Heuvel-Panhuizen, 2003). A similar tension also appears in a more moderate guise in programmes designed to test the relative efficacy of learning via concrete or abstract representations (De Bock, Deprez, Van Dooren, Roelens, & Verschaffel, 2011; Kaminski, Sloutsky, & Heckler, 2008). If one of these approaches were truly superior in a straightforward way then we would surely by now have conclusive evidence of this, so it seems likely that progress will be made by asking not about the extremes of the dichotomy but about the specific implementation and combination of different types of learning experience. As indicated by

the questions in this section, there is much potential for developing and testing specific instructional sequences.

## **E. Designing Effective Interventions**

19. Which domain-specific foundational competencies are most malleable and when in developmental time? And does their malleability impact on other aspects of mathematical performance?
20. What are the features (including content of intervention or instruction and characteristics of children) of current successful interventions and instruction?
21. What concepts and skills need to be in place for success in algebra, how are these treated in different curricula, and what learning outcomes result if these treatments and their sequencing are experimentally compared?
22. What are the most effective interventions for children with dyscalculia, when is the best time for intervention, and which factors best predict the response to intervention?

These questions address issues already raised, but in a more practical way. They focus on the nuances of applied settings, recognising that the effectiveness of any given intervention might depend upon characteristics of the individuals for whom it is intended. The questions also highlight the two-way relationship between theoretical development and experimental testing: by interceding with an intervention based on theoretical propositions about the causes of mathematical errors and difficulties, researchers test both the intervention and the underlying theoretical propositions. It is therefore encouraging that effective interventions do exist and have been tested at scale. Some of these have been designed to allow for individual differences among students who are falling seriously behind, by tailoring research-based interventions to individuals after detailed diagnoses (Holmes & Dowker, 2013; Torgerson et al., 2013). Others have been designed to change instruction in ordinary classrooms by reducing unhelpful regularities in teaching materials and by deliberately designing activities that have demonstrable positive effects (McNeil, Fyfe, & Dunwiddie, 2015). Such interventions so far have been primarily directed at comparatively young

children, which is natural due to levels of research-based knowledge and to serious concerns about the cultural and economic implications of children falling behind at early educational stages. But the forms of such interventions, when combined with localised knowledge about typical understandings of a broader range of concepts, have potential for improving instruction at higher levels too.

## **F. Developing Valid and Reliable Measures**

23. Are there reliable and valid methods for measuring ANS acuity that can be used to track development over an extended period of time (infancy – childhood – adulthood)?
24. How can we measure informal numeracy experiences validly and reliably?
25. How can we measure the variability in children's mathematical experiences outside of school?
26. How can we develop reliable and valid measures of understanding of key mathematical concepts?

These questions capture issues that pervade the entire list: to obtain convincing evidence within and across these questions, we must give appropriate attention to our measures of key predictors and outcomes. In some parts of the literature, debates about these issues are both obvious and detailed. For example, in the substantial body of work on the ANS, methods have become an overt subject of debate as comparisons across research studies reveal serious inconsistencies (Clayton, Gilmore, & Inglis, 2015; De Smedt et al., 2013). In other areas, the discussion is less specific: for instance, methods have been advanced for measuring constructs such as understanding in relation to certain concepts, but there is a lack of agreement over what constitutes evidence of conceptual understanding (Crooks & Alibali, 2014). At higher levels, with exceptions such as the Calculus Concept Inventory (Epstein, 2013), there are few posited standard measures of mathematical knowledge and understanding. The listed questions reflect the need for valid and reliable measures of

understanding and performance across all levels of mathematics. Attention to multiple methods, to their details and differences, is certainly a positive step.

## Critical Reviews

In this section we extend the preceding commentaries for those questions from two sections: Section A (the nature of mathematical thinking) and Section B (developmental trajectories and their interactions).

### Elucidating the Nature of Mathematical Thinking

Recall that the questions in this section are:

1. Do infants really have a 'sense of number' or are they merely sensitive to quantitative dimensions?
2. How do feelings of correctness or doubt arise when people are doing mathematics, and how do they influence mathematical reasoning and response to instruction?
3. What strategies do experts and students use when evaluating whether a mathematical assertion is true or false?
4. What mathematical cognition do experts use in their practice and what methods can researchers use to provide an accurate portrayal of this?

By their nature these questions present serious methodological challenges. A wealth of data suggests that infants have an innate ability to process quantitative information (Izard, San, Spelke & Streri, 2009); habituation and preferential-looking studies indicate that infants can discriminate between sets of small (Starkey & Cooper, 1980) or large quantities (Xu & Spelke, 2000; but see Clearfield & Mix, 2001, for an alternative perspective). It has been argued that these numerical discrimination abilities are guided by two separate systems, *Parallel Individuation* for small quantities and the ANS for large quantities (Feigenson et al., 2004); both systems are believed to be multi-modal in the sense that they aid the discrimination of visual, auditory and action-based quantitative stimuli (Carey, 2009; Lipton &

Spelke, 2004). But a variety of additional characteristics can vary with quantity: for visual stimuli, for instance, increases in quantity generally co-occur with increases in surface area or contour length (Mix & Cheng, 2012). Although researchers may attempt to control for these extraneous perceptual variables, they cannot all be controlled simultaneously (see Cantrell & Smith, 2013 for an extensive review). Therefore, previous findings may be interpreted as indicating that infants are sensitive to changes in a variety of general magnitude dimensions, such as duration, length, convex hull, and area, in addition to quantity (Gebuis & Reynvoet, 2012; Newcombe, Levine & Mix, 2015). The question of which is the most accurate interpretation has yet to be resolved.

Regarding more complex mathematics we know even less. This is important because in mathematics education there is a large movement for students to engage in authentic mathematical practices – to learn to approach problems in a manner epistemologically consistent with the thinking of professional mathematicians (e.g., CCCSI, 2012; Lampert, 1990; NCTM, 2000; Sfard, 1998). Designing for such learning requires an understanding of how mathematical experts complete mathematical tasks and how professional mathematicians practise their craft.

In particular, a desired outcome is for students to learn to solve non-routine problems (e.g., CCCSI, 2012; NCTM, 2000; Schoenfeld, 1985), a process that often involves making conjectures and estimating their veracity. Such activity likely involves many behaviors and competences, one of which is constructing and interpreting mathematical representations. Both analysis of expert protocols (Koedinger & Anderson, 1990; Schoenfeld, 1985; Weber & Alcock, 2004) and mathematicians' reflections on their own practice (Hadamard, 1945; Polya, 1957) indicate that mathematicians' judgments about truth values often involve diagrams, and that their use of these can interact with symbolic reasoning in complex ways (e.g., Koedinger & Anderson, 1990). But there is a growing body of literature indicating that students' propensities to use diagrams have little correlation with their mathematical



achievement (Presmeg, 2006), and we do not yet know why students do not typically reap the same benefits from multiple representations as do mathematicians.

We also lack a clear understanding of what evidence mathematicians use to gain certainty (or high degrees of confidence) in mathematical conjectures. The traditional claim is that mathematicians rely on proofs – logically rigorous deductive arguments (e.g., Griffiths, 2000). But this only raises questions about how such proofs are produced and how mathematicians judge whether they are valid. Given that students usually lack proficiency in constructing and evaluating deductive proofs (e.g., Healy & Hoyles, 2000), even after completing curricula explicitly designed to teach these skills, better models of the proving process could have important implications for mathematics instruction.

Even in the absence of proof, mathematicians must still sometimes make subjective estimations about the plausibility of conjectures. Franklin (2013) argued that such judgments partially determine what problems a mathematician chooses to investigate, and Devlin (2002) proposed that the way mathematicians check arguments for correctness is determined by how much confidence or doubt they have in those arguments. However, there is almost no systematic research into how mathematicians do, or students should, make such judgments. Clearly mathematicians and students do experience varying levels of confidence and doubt about mathematical claims and arguments, so it is natural to ask how these arise and how they influence reasoning and behavior; these are open questions.

We conclude this section by acknowledging the methodological difficulties inherent in investigating students' and mathematicians' mathematical practices. Developing conjectures and forming and evaluating arguments are time-consuming tasks that rely extensively on an individual's background knowledge. Experts do not necessarily address these tasks in the same ways, and the obvious approaches to investigating their practices each have limitations. An investigator can ask individuals to describe their own reasoning processes, and many researchers have employed verbal protocol analyses in attempts to model

students' and mathematicians' problem solving (e.g., Schoenfeld, 1985). But we know that individuals' self-reports about their own reasoning are often inaccurate (e.g., Nisbett & Wilson, 1977). An alternative is to conduct a naturalistic study, but much mathematical reasoning is done in isolation and silence, so this might provide only limited information. Moreover, a naturalistic study incurs substantial time costs: even in school environments, developing and testing a single conjecture can take a class period or longer, and in expert practice, creating and testing a conjecture can span months or years. Naturalistic studies have taken on this challenge, observing mathematicians in meetings over several months (e.g., Greiffenhagen & Sharrock, 2011; Smith, 2012). But they necessarily have relatively small numbers of participants, making it difficult to generalize the findings to the broader population of mathematicians, especially those who work in different domains. And, of course, analyzing the data requires the investigators to be at least partially conversant with cutting-edge mathematics.

One more tractable methodological alternative is to present individuals with proxy tasks, simplified versions of the activities that we would like to investigate that can be completed in a relatively short time and that do not require extensive background knowledge. Using proxy tasks does not guarantee that behavioral patterns observed would be present in authentic situations, but such tasks do permit the investigator to competently analyze the data, and they are beginning to enable progress at least in confirming or debunking common claims. Recent eye-movement studies, for instance, have confirmed that undergraduate students pay less attention than mathematicians to words (as compared with symbols) in purported proofs, and that they are less likely to shift their attention around in a manner consistent with seeking deductive justifications (Inglis & Alcock, 2012). Other recent studies have revealed that mathematicians disagree on the validity of even fairly simple mathematical arguments (Inglis, Mejía-Ramos, Weber, & Alcock, 2013), that they respond in more nuanced ways than is traditionally believed to arguments based on diagrams (Inglis & Mejía-Ramos, 2009a) or presented as from an authoritative source (Inglis & Mejía-Ramos, 2009b), and that their

sense of beauty in a mathematical proof is not systematically related to that proof's perceived simplicity (Inglis & Aberdein, 2014). Several of these findings challenge conventional claims about the nature of mathematical expertise by revealing more than expected heterogeneity in the experiences and judgments of mathematicians. They thus constitute steps toward a more informed view of what mathematics students should learn.

## **Charting Developmental Trajectories and Their Interactions**

The questions in this section are:

1. Are there alternative cognitive pathways by which one can be a successful mathematics student? If yes, what does this imply for typical research methods used by mathematical cognition researchers?
2. How are different mathematical skills (including representing number, counting, performing arithmetic, using fractions) and their developmental trajectories related to each other?
3. How is developmental dyscalculia qualitatively different from arithmetic performance at the lower end of the normal distribution?
4. As we build causal models of pathways to mathematics learning and thinking, what models emerge for compensatory mechanisms – cognitive skills or abilities that allow individuals to overcome shortcomings in essential skills, and that contribute to patterns of behaviors?

In comparison with the questions discussed in the previous section, research pertaining to these is comparatively advanced. But it is by no means conclusive: the available data suggest that multiple potential pathways to mathematical success exist, but do not specify precisely what those pathways are. This is evident among children who have mathematical difficulties of known (e.g., Dennis et al., 2009) or unknown etiology (e.g., Geary, 2011, Kaufmann et al., 2013), the most easily differentiated of whom have readily classifiable developmental diagnoses such as fragile X syndrome, spina bifida, and 22q deletion

syndrome (De Smedt, Swillen, Verschaffel, & Ghesquière, 2009; Mazzocco, Quintero, Murphy, & McCloskey, 2016). Some group differences, while disorder-specific, may provide models for variation in the general public if origins for and mechanisms underlying different pathways and trajectories were well understood. This makes inquiry into these trajectories particularly valuable.

To understand differences in developmental trajectories, researchers attend to relative deficits and assets – a neuropsychological approach that reveals potential dissociations of mathematics skills and that thus extends knowledge of basic cognitive processes and potential routes to both dyscalculia and typical mathematical achievement. For instance, school-age children with 22q deletion syndrome (22qDS) have intact fact retrieval skills but relatively impaired enumeration of quantities (see the review by de Smedt et al., 2009); they also exhibit intact word learning, but under-count enumeration errors for sets exceeding four (Quintero, Beaton, Harvey, Ross, & Simon, 2014). Whether these enumeration errors reflect difficulties with mapping of numbers to quantities (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003), with spatial attention (Simon, Bearden, Mc-Ginn, & Zackai, 2005), or with other basic processes has important implications for the origins of developmental trajectories. But, in all possible cases, their existence demonstrates a dissociation between components of mathematical cognition.

Similar dissociations are evident across other conditions. For instance, mathematical performance remains comparatively unimpaired in children with 22qDS when tapping the verbal system (Mazzocco et al., 2016), but not when engaging in nonverbal tasks. In contrast, whereas children with spina bifida meningomyelocele show delayed counting principles in early childhood and poor numeracy across the lifespan, their mathematical fact retrieval performance is not correlated with presence or absence of a learning disability (English, Barnes, Taylor, & Landry, 2009) but rather with level of fine motor skills (Barnes et al., 2006). In children with Williams syndrome, the dissociation between numerical skills is not only in nature but also in timing: abstract representations of small quantity are age

appropriate during infancy, as is the ability to track small item sets. But these skills do not remain intact throughout life, and impaired approximate number representations of large sets emerge even in early childhood. Remarkably, despite this profile, young children with Williams Syndrome (WS) outperform other young girls with fragile X syndrome on counting principles such as cardinality.

These systematic profile analyses are each guided by theories linked to specific biochemical, neuroanatomical, or neuropsychological characteristics of known conditions (as reviewed by Dennis et al., 2009). As such, principled approaches allow us to test theories of the origins of mathematics learning trajectories, or to ask when pathways diverge or become unique. Possible factors in divergence include impaired motor function that prevents motor exploration in infants and toddlers with spina bifida (and that limits finger use in older children), and disrupted development of the visuospatial processing system in 22qDS or WS. But these factors are not syndrome specific, so the mathematical assets observed in these and other populations have implications for potential compensatory mechanisms – routes via which we might promote mathematical learning in any children who share assets in these domains.

This argument about assets extends to domain-general skills like working memory (English et al., 2009), and of course developmental differences are also observed in children with typical mathematics achievement (LeFevre et al., 2010). To date, studies of typical populations have focused on numerical skills, fractions, fact retrieval/word problem solving, or general mathematical achievement, depending on the researcher's discipline. Measured potential mediators or moderators frequently include executive functions (e.g., working memory, inhibitory control; Bull, Espy, & Wiebe, 2008), processing speed, and, more recently, indices of mathematics anxiety or motivational factors (Beilock et al., 2010; Wang et al., 2015). Importantly, most of the relationships to emerge from these studies are nonlinear. Outcomes may interact with problem type (Agostino, Johnson, & Pascual-Leone, 2010), stage of development (Ansari, 2010), motivation or mathematics anxiety (Beilock et al., 2010;

Wang et al., 2015), and other known influences (e.g., reading disability, stereotype threat, motor function) work directly on mathematical outcomes or indirectly moderate known associations (for instance, children with a reading disability or moderate math anxiety might be more or less responsive to an intervention).

Ideally, we would draw comparisons across studies that ask similar questions or address different disorders; this would help to identify common dissociations, alternative compensatory paths, and indicators of whether children are good candidates for specific compensation-based interventions. But, at present, an explicit comparison of this sort is problematic. At minimum it requires commonality in at least some of the measures used to assess skills (mathematical or otherwise), overlap in ages assessed, and similar methodological approaches to study design and data analysis. This returns us to the overarching questions about methodology raised in Section F: better consistency in design and reporting would contribute to both disorder-specific knowledge and knowledge of mathematical cognition in the broadest sense.

## Discussion

The exercise reported in this paper led to a list of questions that reflects a broad approach to understanding human mathematical cognition, encompassing the effects of environmental and affective factors, of domain-general and domain-specific cognitive factors, and of teaching interventions. It covers mathematical cognition as it does or can develop from infancy to expertise: from basic representations of number in the brain, through informal and formal learning and ways in which these might be studied and supported, to the thinking of expert mathematicians. This variation, however, should not be seen as overwhelming: the list exhibits an encouraging level of coherence among the main concerns of researchers at all mathematical levels and from all of the disciplines involved in our exercise. Indeed, although the conference was characterised by a spirit of openness, engagement, and a willingness to do the work necessary to iron out any misunderstandings that arose due to

different terminologies, this was hardly required – genuine disagreements were few, which encourages us to believe that our different approaches are converging on similar conceptions of the key issues.

That said, any exercise like this is naturally limited by the experience and knowledge of its participants, and must trade off breadth of representation against depth of focus. With a comparatively small exercise like this, we leant towards depth: all of our participants work on mathematical cognition as a major strand of their own research. One consequence was that minority interests are less well reflected in our list of questions; in particular, although questions related to cognitive neuroscience arose in the drafting stages (see [https://figshare.com/articles/S1\\_Dataset\\_xlsx/1299358](https://figshare.com/articles/S1_Dataset_xlsx/1299358)), they did not survive the winnowing process. This could be seen as reflecting a primary focus among our participants on questions of fairly direct concern to educators; fundamental questions about representations in the brain are perhaps, at present, a step further removed from these concerns. Similarly, resource restrictions and the UK location of the funder skewed the participation toward local participants and toward those who publish primarily in English. We fully acknowledge that other researchers – especially those with broader interests, or those working primarily in other regions or languages – might take different views, and we hope that one function of this paper will be to stimulate debate.

We hope also to stimulate debate on prioritisation of research endeavours, toward which end we should clarify that we intend this paper as a resource, not a manifesto: we do not wish to make actual or speculative value judgments about how our fellow researchers use their time and resources. Some bodies do make such judgments: funders are entitled to establish agendas and research priorities. But we consider it inappropriate to use our list – whether pre-emptively or retrospectively – as a basis for judging the worthiness of others' work. We nevertheless believe that the spirit of academic freedom would be preserved if an outcome of our exercise is a set of important shared considerations that becomes refined by greater interdisciplinary communication. Accordingly we hope that others will consider addressing

some questions directly, and relating their ongoing work to a broader swathe of research pertinent to mathematical cognition.

We are also aware that our list of questions raises issues about practicality: some questions could be addressed via small research projects run by small teams, but some would necessitate larger-scale, multi-lab investigations. The list also raises natural questions about sequencing; for instance, it may seem that question 25 (*How can we measure the variability in children's mathematical experiences outside of school?*) must be answered before question 7 (including *What are reliable early and later longitudinal predictors of the development of number skills, arithmetic and other aspects of mathematics?*). But this presupposes that a perfect measure is possible, and that it is possible to know in isolation that we have found it. In fact, such a measure must necessarily be constructed in conjunction with research about the influence of various factors on development: there is no point in measuring perfectly a quantity that turns out not to have influence in the world (arguably, the abstract nature of theoretical constructs means that this is not even possible). There are also more prosaic concerns: we could invest in developing excellent measures for constructs that we currently believe to be important, but what if they turn out not to be? Without some concurrent effort to establish at a rough level whether outside-of-school experiences have as much influence as, say, the first year in school, that investment could be wasted (at least from the perspective of educational development, if not from that of basic research).

So, although our questions represent an informed, collective, current perspective, it would be folly to set ourselves up as able to make predictions about what is destined to become the really important breakthroughs or insights. These will come about due to continued reflection informed by new developments across our fields, so we hope – as stated in the introduction – to promote communication and interdisciplinary work across the fields of mathematics education, psychology and cognitive neuroscience by drawing attention to what is common in their methods, theories, and ultimate research goals.



## Funding/Financial Support

The reported exercise was part-funded by a Royal Society International Conference Grant, along with funding from Economic and Social Research Council grants (ES/L010089/1 and RES-062-23-3280), the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, the Canada Research Chairs (CRC) Program, a Royal Society Worshipful Company of Actuaries Research Fellowship, a Royal Society Shuttleworth Educational Research Fellowship, and a Royal Society Dorothy Hodgkin Research Fellowship.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Competing Interests

The authors have declared that no competing interests exist.

## References

- Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology*, 105, 286-305. doi:10.1016/j.jecp.2009.09.006
- Alcock, L., Gilmore, C., & Inglis, M. (2013). Experimental methods in mathematics education research. *Research in Mathematics Education*, 15, 97-99. doi:10.1080/14794802.2013.797731

- Ansari, D. (2010). Neurocognitive approaches to developmental disorders of numerical and mathematical cognition: The perils of neglecting the role of development. *Learning and Individual Differences*, 20, 123-129. doi:10.1016/j.lindif.2009.06.001
- Barnes, M. A., Wilkinson, M., Khemani, E., Boudesquie, A., Dennis, M., & Fletcher, J. M. (2006). Arithmetic processing in children with spina bifida calculation accuracy, strategy use, and fact retrieval fluency. *Journal of Learning Disabilities*, 39, 174-187. doi:10.1177/00222194060390020601
- Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, 107, 1860-1863. doi: 10.1073/pnas.0910967107
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, 33, 205-228. doi:10.1080/87565640801982312
- Cantrell, L. & Smith, L.B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, 128, 331-352. doi:10.1016/j.cognition.2013.04.008
- Carey S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.
- Clayton, S., Gilmore, C. & Inglis, M. (2015). Dot comparison stimuli are not all alike: The effect of different visual controls on ANS measurement. *Acta Psychologica*, 161, 177-184. doi:10.1016/j.actpsy.2015.09.007
- Clearfield, M. W., & Mix, K. S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets. *Journal of Cognition and Development*, 2, 243-260. doi:10.1207/S15327647JCD0203\_1
- Clements, D. H., & Sarama, J. (2004). Learning Trajectories in Mathematics Education. *Mathematical Thinking and Learning*, 6, 81-89. doi:10.1207/s15327833mtl0602\_1

- Cohen Kadosh, R., Dowker, A., Heine, A., Kaufmann, L., & Kucian, K. (2013). Interventions for improving numerical abilities: Present and future. *Trends in Neuroscience and Education*, 2, 85–93. doi:10.1016/j.tine.2013.04.001
- Common Core Content Standards Initiative (2012). Common core state standards initiative: Mathematical standards. Retrieved from [www.corestandards.org/wp-content/uploads/Math\\_Standards1.pdf](http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf)
- Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, 34, 344–377. doi:10.1016/j.dr.2014.10.0011
- De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or Concrete Examples in Learning Mathematics? A Replication and Elaboration of Kaminski, Sloutsky, and Heckler's Study. *Journal for Research in Mathematics Education*, 42(2), 109–126. Retrieved from [www.jstor.org/journal/jresematheduc](http://www.jstor.org/journal/jresematheduc)
- De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48–55. doi:10.1016/j.tine.2013.06.001
- De Smedt, B., Swillen, A., Verschaffel, L., & Ghesquière, P. (2009). Mathematical disabilities in children with 22q11.2 deletion syndrome: A review. *Developmental Disabilities Research Reviews*, 15, 4-10. doi:10.1002/ddrr.44
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20, 487-506. doi:10.1080/02643290244000239
- Dennis, M., Berch, D. B., & Mazzocco, M. M. M. (2009). Mathematical learning disabilities in special populations: Phenotypic variation and cross-disorder comparisons. *Developmental Disabilities Research Reviews*, 15, 80-89. doi:10.1002/ddrr.54

- Devlin, K. (2002). The Millenium Problems: The seven greatest unsolved mathematical puzzles of our time. New York, NY: Basic Books.
- English, L. H., Barnes, M. A., Taylor, H. B., & Landry, S. H. (2009). Mathematical development in spina bifida. *Developmental disabilities research reviews*, 15, 28-34. doi:10.1002/ddrr.48
- Epstein, J. (2013). The Calculus Concept Inventory - Measurement of the effect of methodology in mathematics. *Notices of the American Mathematical Society*, 60, 1018–1026. doi:10.1090/noti1033
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in cognitive sciences*, 8, 307-314. doi:10.1016/j.tics.2004.05.002
- Franklin, J. (2013). Non-deductive logic in mathematics: The probability of conjectures. In A. Aberdein & I. Dove (Eds.) *The argument of mathematics* (pp. 11-30). Springer: Dordrecht.
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., ... Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology*, 46, 1731–1746. doi:10.1037/a0020662.
- Geary, D. C (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Educating the evolved mind* (Vol. 2, pp. 1-99). Greenwich, CT: Information Age.
- Geary, D. C. (2010). Missouri longitudinal study of mathematical development and disability. In V. Connelly, A. L. Barnett, J. E. Dockrell & A. Tolmie (Eds), *Understanding number development and difficulties* (pp. 31–49). doi:10.1348/97818543370009X12583699332410
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, 47, 1539–1552. doi:10.1037/a0025510

- Gebuis, T. & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology, General*, 141, 642-648. doi:10.1037/a0026218
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS ONE*, 8(6), e67374. doi:10.1371/journal.pone.0067374
- Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science*, 25, 789-798. doi:10.1177/0956797613516471
- Goldin-Meadow, S., Wagner Cook, S., & Mitchell, Z.A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20, 267-272. doi: 10.1111/j.1467-9280.2009.02297.x
- Gravemeijer, K. P. E. (1994). *Developing realistic mathematics education*. Utrecht: Freudenthal Institute.
- Greiffenhagen, C., & Sharrock, W. (2011). Does mathematics look certain in the front, but fallible in the back? *Social Studies of Science*, 41, 839-866. doi:10.1177/0306312711424789
- Griffiths, P.A. (2000). Mathematics at the turn of the millennium. *American Mathematical Monthly*, 107, 1-14. doi:10.2307/2589372
- Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Princeton, NJ: Princeton University Press.
- Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 34, 396-428. doi:10.2307/749651
- Heuvel-Panhuizen, M. V. D. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics*, 54, 9-35. doi:10.1023/B:EDUC.0000005212.03219.dc

- Holmes, W., & Dowker, A. (2013). Catch Up Numeracy: a targeted intervention for children who are low-attaining in mathematics. *Research in Mathematics Education*, 15, 249–265. doi:10.1080/14794802.2013.803779
- Hsu, C.-C. & Sandford, B.A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1-8. Retrieved from pareonline.net
- Inglis, M. & Aberdein, A. (2014). Beauty is not simplicity: An analysis of mathematicians' proof appraisals. *Philosophia Mathematica*, 23, 87-109. doi:10.1093/phimat/nku014
- Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43, 358–390. doi:10.5951/jresmetheduc.43.4.0358
- Inglis, M. & Mejía-Ramos, J. P. (2009a). On the persuasiveness of visual arguments in mathematics. *Foundations of Science*, 14, 97-110. doi:10.1007/s10699-008-9149-4
- Inglis, M., & Mejía-Ramos, J. P. (2009b). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25–50. doi:10.1080/07370000802584513
- Inglis, M., Mejía-Ramos, J.P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5, 270-282. doi:10.1111/tops.12019
- Izard, V., Sann, C., Spelke, E. & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, 106, 10382–10385. doi:10.1073/pnas.0812142106
- Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: Two components of a relational conception of the equals sign. *Journal of Experimental Child Psychology*, 113, 166-176. doi:10.1016/j.jecp.2012.05.003
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, 320, 454–455. doi:10.1126/science.1154659

- Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Gobel, S. M., Grabner, R. H., ...  
 Nuerk, H.-C. (2013). Dyscalculia from a developmental and differential perspective.  
*Frontiers in Psychology*, 4, 1-5. doi:10.3389/fpsyg.2013.00516
- Klibanoff, R.S., Levine, S.C., Huttenlocher, J., Vasilyeva, M., & Hedges, L.V. (2006).  
 Preschool children's mathematical knowledge: The effect of teacher "math talk".  
*Developmental Psychology*, 42, 59-69. doi:10.1037/0012-1649.42.1.59
- Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks:  
 Elements of expertise in geometry. *Cognitive Science*, 14, 511-550.  
 doi:10.1207/s15516709cog1404\_2
- Lampert, M. (1990). When the problem is not the question and the solution is not the answer:  
 Mathematical knowing and teaching. *American Educational Research Journal*, 27,  
 29-63. doi:10.3102/00028312027001029
- LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., &  
 Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of  
 performance. *Child development*, 81, 1753-1767. doi:10.1111/j.1467-  
 8624.2010.01508.x
- LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J.  
 (2009). Home numeracy experiences and children's math performance in the early  
 school years. *Canadian Journal of Behavioural Science*, 41, 55-66.  
 doi:10.1037/a0014532
- Lipton, J. & Spelke, E. (2004). Discrimination of large and small numerosities by human  
 infants. *Infancy*, 5, 271-290. doi:10.1207/s15327078in0503\_2
- Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate  
 number system underlies mathematical learning disability (dyscalculia). *Child  
 development*, 82, 1224-1237. doi:10.1111/j.1467-8624.2011.01608.x

- Mazzocco M.M., Quintero, A., Murphy, M.M., & McCloskey M. (2016). Genetic syndromes as model pathways to mathematical learning difficulties: Fragile X, Turner, and 22q deletion syndromes. In D. B. Berch, D.C. Geary, & K. Mann-Koepeke (Eds), *Mathematical Cognition and Learning Serie: Neural Substrates and Genetic Influences* (Vol. 2, pp.325-359). London, United Kingdom: Elsevier.
- Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: developmental and educational implications. *Advances in child development and behavior*, 42, 197-243. doi:10.1016/B978-0-12-394388-0.00006-X
- McNeil, N. M. (2014). A change-resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, 8, 42–47. doi:10.1111/cdep.12062
- McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology*, 107, 423-436. doi:10.1037/a0037687
- National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.
- National Research Council. (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Newcombe, N., Levine, S. & Mix, K. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs Cognitive Science*, 6, 491-505. doi:10.1002/wcs.1369
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84, 231-259. doi:10.1037/0033-295X.84.3.231
- Polya, G. (1957). *How to solve it: A new aspect of mathematical method*. London: Penguin.



- Prediger, S., Gravemeijer, K., & Confrey, J. (2015). Design research with a focus on learning processes: An overview on achievements and challenges. *ZDM Mathematics Education*, 47, 877-891. doi:10.1007/s11858-015-0722-3
- Presmeg, N. C. (2006). Research on visualization in learning and teaching mathematics. In A. Gutierrez & P. Boero (Eds.) *Handbook of research on the psychology of mathematics education*, (pp. 205-235). Sense Publishers: Rotterdam.
- Quintero, A. I., Beaton, E. A., Harvey, D. J., Ross, J. L., & Simon, T. J. (2014). Common and specific impairments in attention functioning in girls with chromosome 22q11.2 deletion, fragile X or Turner syndromes. *Journal of Neurodevelopmental Disorders*, 6(5), 1-15. doi 10.1186/1866-1955-6-5
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, 20, 110–122. doi:10.1016/j.lindif.2009.10.005
- Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. Cohen Kadosh & A. Dowker (Eds.), *Oxford handbook of numerical cognition* (pp.1118-1134). Oxford, UK: Oxford University Press.
- Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.
- Schoenfeld, A. H. (2004). The Math Wars. *Educational Policy*, 18, 253–286. doi:10.1177/0895904803260042
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13. doi:10.3102/0013189X027002004
- Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, 59, 833-851. doi:10.2307/1130252
- Simon, T. J., Bearden, C. E., Mc-Ginn, D. M., & Zackai, E. H. (2005). Visuospatial and numerical cognitive deficits in children with chromosome 22q11.2 deletion syndrome. *Cortex*, 41, 145-155. doi:10.1016/S0010-9452(08)70889-X

- Skemp, R. R. (1976). Relational Understanding and Instrumental Understanding. *Mathematics Teaching*, 77, 20–27. Retrieved from [www.atm.org.uk](http://www.atm.org.uk)
- Smith, M. (2012). *Methods of mathematical struggle*. Unpublished doctoral dissertation: San Diego State University.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, 36, 404–411. doi:10.2307/30034943
- Starkey, P., & Cooper, R. G. Jr. (1980) Perception of numbers by human infants. *Sciences*, 210, 1033-1035. doi:10.1126/science. 7434014
- Sutherland, W. J., Armstrong-Brown, S., Armsworth, P. R., Tom, B., Brickland, J., Campbell, C. D., ... Watkinson, A. R. (2006). The identification of 100 ecological questions of high policy relevance in the UK. *Journal of Applied Ecology*, 43, 617–627. doi:10.1111/j.1365-2664.2006.01188.x
- Sutherland, W. J., Bellingan, L., Bellingham, J. R., Blackstock, J. J., Bloomfield, R. M., Bravo, M., ... Zimmern, R. L. (2012). A collaboratively-derived science-policy research agenda. *PLoS ONE*, 7(3), e31824. doi:10.1371/journal.pone.0031824
- Sutherland, W. J., Fleishman, E., Mascia, M. B., Pretty, J., & Rudd, M. A. (2011). Methods for collaboratively identifying research priorities and emerging issues in science and policy. *Methods in Ecology and Evolution*, 2, 238–247. doi:10.1111/j.2041-210X.2010.00083.x
- Torgerson, C., Wiggins, A., Torgerson, D., Ainsworth, H., & Hewitt, C. (2013). Every child counts: Testing policy effectiveness using a randomised controlled trial, designed, conducted and reported to CONSORT standards. *Research in Mathematics Education*, 15, 141–153. doi:10.1080/14794802.2013.797746
- Van Hoof, J., Lijnen, T., Verschaffel, L., & Van Dooren, W. (2013). Are secondary school students still hampered by the natural number bias? A reaction time study on fraction comparison tasks. *Research in Mathematics Education*, 15, 154–164. doi:10.1080/14794802.2013.797747

- Wang, Z. Lukowski, S.L., Hart, S.A., Lyons, I. M., Thompson, L.A., Kovas, Y., Mazzocco, M. M., Plomin, R. & Petrill, S. A. (2015). Is mathematical anxiety always bad for math learning? The role of math motivation. *Psychological Science*, 26, 1863-1876. doi:10.1177/0956797615602471
- Weber, K. & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56, 209-234. doi:10.1023/B:EDUC.0000040410.57253.a1
- Weber, K., Inglis, M., & Mejía-Ramos, J. P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, 46, 36–58. doi:10.1080/00461520.2013.865527
- Wu, S. S., Meyer, M. L., Maeda, U., Salimpoor, V., Tomiyama, S., Geary, D. C., & Menon, V. (2008). Standardized assessment of strategy use and working memory in early mental arithmetic performance. *Developmental neuropsychology*, 33, 365-393. doi:10.1080/87565640801982445
- Xu, F. & Spelke, E. (2000). Large number discrimination in 6-month old infants. *Cognition*, 74, B1–B11. doi:10.1016/S0010-0277(99)00066-9.

## Endnotes

<sup>i</sup> The journals and conferences in this section are not intended to be exhaustive or typical for all researchers within a specific discipline; instead they serve to illustrate the diverse publication outlets within the field of mathematical cognition.

<sup>ii</sup> These exercises have typically been influential, with the resulting papers attracting large numbers of citations both in academic research and in policy documents.

<sup>iii</sup> A conference that took place at the Kavli Royal Society International Centre in August 2014.

ACCEPTED MANUSCRIPT